

Characterization and cloning of fasciclin I and fasciclin II glycoproteins in the grasshopper

(membrane proteins/immunoaffinity purification/neuronal recognition/growth cone guidance)

PETER M. SNOW^{*†}, KAI ZINN^{*†}, ALLAN L. HARRELSON^{*†}, LINDA MCALLISTER^{*†}, JIM SCHILLING[‡],
MICHAEL J. BASTIANI^{*§}, GEORGE MAKK[¶], AND COREY S. GOODMAN^{*†||}

^{*}Department of Biological Sciences, Stanford University, Stanford, CA 94305; [‡]California Biotechnology, Mountain View, CA 94043; and [¶]Department of Psychiatry, Stanford University Medical School, Stanford, CA 94305

Communicated by Gerald M. Rubin, March 9, 1988

ABSTRACT Monoclonal antibodies were previously used to identify two glycoproteins, called fasciclin I and II (70 and 95 kDa, respectively), which are expressed on different subsets of axon fascicles in the grasshopper (*Schistocerca americana*) embryo. Here the monoclonal antibodies were used to purify these two membrane-associated glycoproteins for further characterization. Fasciclin II appears to be an integral membrane protein, whereas fasciclin I is an extrinsic membrane protein. The amino acid sequences of the amino terminus and fragments of both proteins were determined. Using synthetic oligonucleotide probes and antibody screening, we isolated genomic and cDNA clones. Partial DNA sequences of these clones indicate that they encode fasciclins I and II.

Previous studies gave rise to the labeled pathways hypothesis, which predicts that axon fascicles in the embryonic neuropil are differentially labeled by surface recognition molecules used for growth cone guidance (e.g., see refs. 1–6). To identify candidates for axonal recognition molecules, monoclonal antibodies (mAbs) were generated that recognize surface antigens expressed on subsets of axon fascicles in both grasshopper (*Schistocerca americana*) (7) and *Drosophila* (8) embryos.

These mAbs were used to characterize three different membrane-associated glycoproteins, called fasciclin I and II in the grasshopper and fasciclin III in *Drosophila*, which have several features in common. All three proteins (i) are expressed on different subsets of axon fascicles during development, (ii) are regionally expressed on particular portions of embryonic neurons where their axons fasciculate together, (iii) are dynamically expressed during axon outgrowth, and (iv) are expressed outside the developing nervous system at other times and places.

To begin a molecular genetic analysis of the structure and function of these proteins, the genes encoding all three have been cloned. The cloning of fasciclin III from *Drosophila* was previously reported (8). In this paper, we report on the further characterization of the fasciclin I and II glycoproteins in the grasshopper, and on the isolation and partial characterization of cDNA clones encoding both molecules.**

MATERIALS AND METHODS

Protein Purification. Fasciclin I and fasciclin II were purified by using an affinity matrix based on staphylococcal protein-A-Sepharose (Pharmacia) as described (7, 9). Preparative NaDodSO₄/polyacrylamide gel electrophoresis (NaDodSO₄/PAGE) and electroelution were performed according to established procedures (10). Removal of excess Co-

massie blue and NaDodSO₄ prior to protein sequence analysis by ion-pair extraction was performed as described by Konigsberg and Henderson (11). For determination of the amino-terminal sequences of intact fasciclin I or cyanogen bromide (CNBr) fragments of fasciclins I and II, electroeluted samples were sequenced on a gas-phase microsequencer (Applied Biosystems, Foster City, CA, model 470A) with an on-line HPLC for analysis of the phenylthiohydantoin of the amino acids, using the manufacturer's standard reagents and programs.

Biochemical Techniques. Treatment of glycoproteins with trifluoromethanesulfonic acid was performed as described by Edge *et al.* (12). Proteins to be cleaved with CNBr were first lyophilized. The samples (1–2 nmol) were solubilized in 200 μ l of 75% (vol/vol) trifluoroacetic acid (Pierce) saturated with CNBr. Cleavage was allowed to proceed for 20–24 hr in the dark at room temperature under nitrogen. Trifluoroacetic acid was removed with a stream of nitrogen and the samples were lyophilized from 1 ml of water three times. The cleaved proteins were subsequently subjected to preparative NaDodSO₄/PAGE and electroelution.

Membranes were prepared from adult grasshopper nervous systems as previously described (7). Membrane proteins were labeled with ¹²⁵I and lactoperoxidase (13). In some experiments, the labeled membranes were treated with 50 mM triethylamine, pH 11.5, for 15 min at 4°C, and the membranes were collected by centrifugation at 12,000 \times g and washed once with the same buffer. The supernatants were combined and subjected to centrifugation at 100,000 \times g for 30 min. The stripped membranes were solubilized in 10 mM triethanolamine, pH 8.2/0.15 M NaCl/1% Nonidet P-40/1 mM phenylmethanesulfonyl fluoride for 30 min at 0°C, and insoluble residue was removed by centrifugation at 100,000 \times g for 30 min. Both supernatant and solubilized membranes were subjected to immunoprecipitation with preformed antibody complexes as described (14). Analytical NaDodSO₄/PAGE was performed according to Laemmli (15). Silver staining was performed as described by Morrissey (16).

Construction of Grasshopper Genomic and cDNA Libraries. The grasshopper genomic library was constructed by standard methods (17) from grasshopper DNA partially digested with *Mbo* I and size-fractionated. The DNA was ligated into *Bam*HI-digested λ J1 (18) vector arms and packaged.

Abbreviation: mAb, monoclonal antibody.

[†]Present address: Department of Biochemistry, University of California, Berkeley, CA 94720.

[§]Present address: Department of Biology, University of Utah, Salt Lake City, UT 84112.

^{||}To whom reprint requests should be sent at [†] address.

**The sequences reported in this paper are being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession nos. J03787, J03788, and J03789 for fasciclins I, II, and III, respectively).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

The grasshopper embryo cDNA libraries were made from poly(A)⁺ RNA from dissected 40–50% developed grasshopper embryos. RNA was purified by lysis in guanidine hydrochloride followed by centrifugation through a cesium chloride cushion (19). Double-stranded cDNA was made by a modification of the protocol of Gubler and Hoffman (20). The cDNA was methylated with *Eco*RI methylase, ligated to *Eco*RI linkers, recut, and size-fractionated on an agarose gel. cDNA larger than 1.8 kilobases (kb) (about 10–20% of the total cDNA molecules) was eluted, ligated into dephosphorylated λ gt11 arms (21), and packaged by using a packaging extract from nonrestricting (*r*[−]) *Escherichia coli* (Stratagene, San Diego, CA). Use of this extract proved to be important, as another aliquot of the same cDNA pool from which nine fasciclin I clones were isolated produced no fasciclin I clones when packaged with an *r*⁺ extract of equal efficiency. The fasciclin I cDNA sequence was found to contain an *Eco* K site.

Oligonucleotide Probes and Library Screens. Oligonucleotide probes were synthesized on an Applied Biosystems synthesizer and 5'-end-labeled with polynucleotide kinase. Hybridizations with the 44-mer probe were done in 6× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) at 42°C, followed by low-stringency 1× SSC washes and a final wash in 3 M tetramethylammonium chloride at 68°C (22). For hybridization to the redundant 20-mer pool, plaques amplified *in situ* (23) were transferred to nylon membranes and hybridized to probe at 10 μ Ci/ml (1 Ci = 37 GBq) in 3 M tetramethylammonium chloride at 51°C. The filters were also washed in 3 M tetramethylammonium chloride (final wash at 59–60°C).

Expression cloning with rat antisera was done by standard methods (8, 21).

Other Molecular Biology Techniques. Southern blotting, subcloning, and other molecular biology techniques were performed by standard methods (24). DNA sequencing was performed on M13 and Bluescript (Stratagene) subclones by the dideoxy method (25). Clones for sequencing were generated from larger clones by unidirectional deletion with exonuclease III (26).

RESULTS

Biochemical Characterization of Fasciclin I and Fasciclin II. Fasciclin I and fasciclin II were initially identified by using mAbs (7). The 3B11 and 8C6 mAbs immunoprecipitate membrane-associated proteins from both adult grasshopper central nervous system (CNS) and 40–50% developed grasshopper embryos of 70 and 95 kDa, respectively. Because these two proteins are expressed on different subsets of axon fascicles during development (Fig. 1), we call them fasciclin I and fasciclin II. Kilogram quantities of solubilized grasshopper embryos were used to purify microgram quantities of each protein by immunoaffinity chromatography (Fig. 2). The purified protein was used for the generation of antisera in rats (Fig. 1), for further biochemical characterization (Fig. 2), or for protein microsequencing (Fig. 4).

Both proteins are glycosylated, as indicated by three different lines of evidence. First, both are bound by the lectin concanavalin A (not shown). Second, the apparent molecular weight of both proteins, when analyzed by NaDodSO₄/PAGE, is significantly decreased by treatment with trifluo-

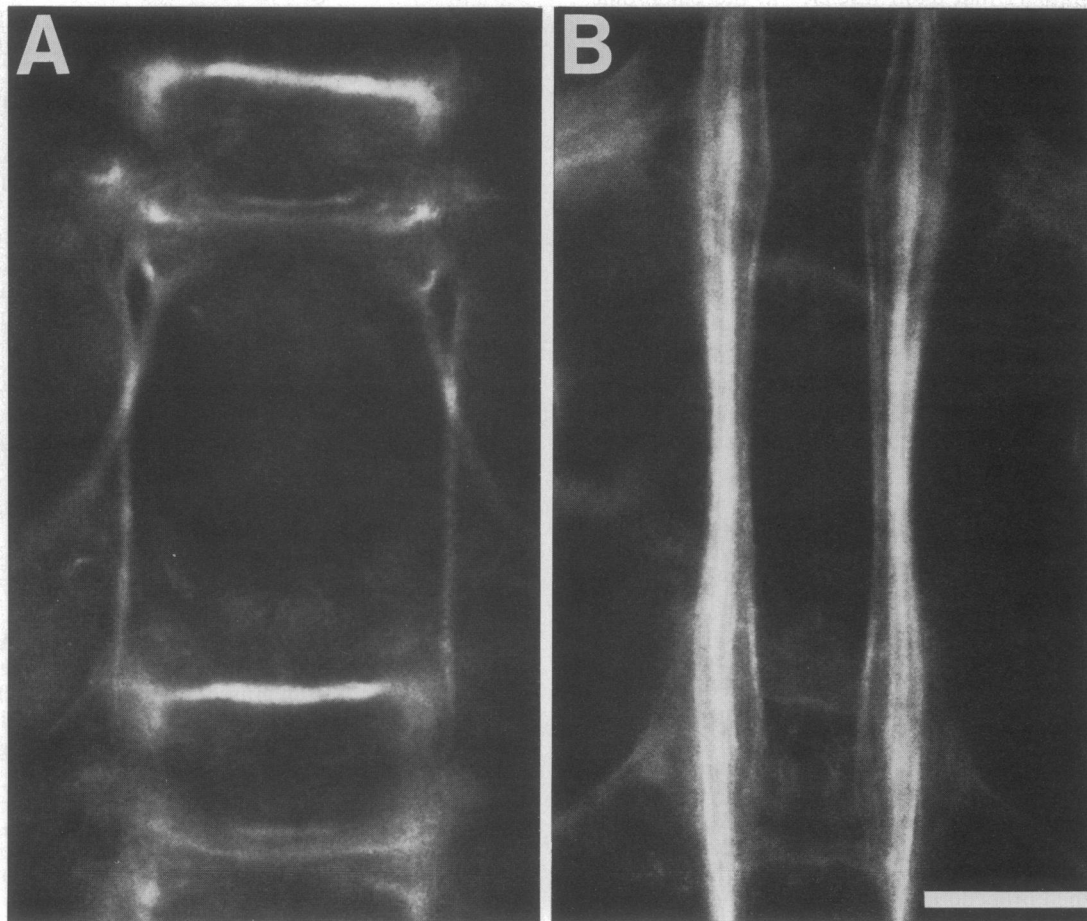


FIG. 1. Expression of fasciclin I (A) and fasciclin II (B) glycoproteins on specific subsets of axon pathways in the grasshopper embryo. Serum antibodies against the two purified proteins recognize the same subsets of axon pathways as do the mAbs against the same proteins, indicating that the proteins are indeed expressed on restricted subsets of axons. (Bar = 50 μ m.)

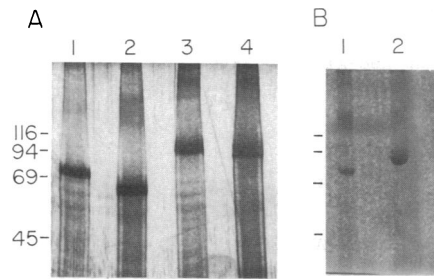


FIG. 2. Biochemical analysis of fasciclin I and fasciclin II. (A) Both fasciclin I and fasciclin II are glycoproteins. Fasciclin I (lanes 1 and 2) and fasciclin II (lanes 3 and 4) were treated with trifluoromethanesulfonic acid (TFMS) and subsequently analyzed on a 10% polyacrylamide gel and visualized by silver staining. Lane 1, fasciclin I, no treatment; lane 2, fasciclin I, TFMS-treated; lane 3, fasciclin II, no treatment; lane 4, fasciclin II, TFMS-treated. (B) Analysis of fasciclins I and II under nonreducing conditions. Affinity-purified fasciclin I (lane 1) and fasciclin II (lane 2) were analyzed on a 10% polyacrylamide gel in the absence of reducing agents and visualized by staining with Coomassie blue. On the left are sizes of markers in kDa.

romethanesulfonic acid (Fig. 2), which has been shown to cleave both N- and O-linked sugars (12). Third, both proteins are immunoprecipitated by the antibody to horseradish peroxidase (27) that recognizes a neural-specific carbohydrate epitope in insects (28). The anti-horseradish peroxidase immunoprecipitates only about 10% of each protein, indicating heterogeneity in the glycosylation of both.

Analysis of the relative mobilities of the native and deglycosylated forms of fasciclin I (Fig. 2A) indicates that the core protein has a molecular mass of 64 kDa with 8 kDa of carbohydrate. Similarly, fasciclin II consists of a polypeptide of 87 kDa and an oligosaccharide component of 6 kDa [this result has been confirmed by the use of a lower percentage acrylamide gel, which allows greater resolution in the appropriate molecular mass range (data not shown)]. The polypeptide components of fasciclin I and fasciclin II do not seem to be related at the level of their primary sequences. Not only are their sizes different, but neither one of them is recognized on immunoblots by the antisera against the other one (data not shown). In addition, two-dimensional peptide maps indicate no similarities in their tryptic fragments (data not shown). Neither fasciclin I nor fasciclin II is covalently linked to itself or to other proteins as indicated by comparison of reduced and nonreduced gels (Fig. 2B). Moreover, no additional proteins are isolated by the antibody affinity columns, indicating that neither protein is tightly associated with other subunits.

While fasciclins I and II are similar in many of their biochemical characteristics, they differ in one respect. Whereas fasciclin II behaves as an integral membrane protein, fasciclin I appears to be an extrinsic membrane protein. Thus, when homogenized adult CNS preparations are subjected to differential centrifugation, both glycoproteins are found to be associated with the membrane and not with the soluble fraction. Further, monoclonal and serum antibodies against both proteins stain the outer surface of axon fascicles in living grasshopper embryos (7).

However, the two proteins behave quite differently when membrane preparations are subjected to alkaline conditions. As shown in Fig. 3 (compare lanes 3 and 4), fasciclin II is refractory to extraction at pH 11.5. While in this experiment the background in the area of fasciclin II in the supernatant lane is high, subsequent immunoprecipitations have confirmed the absence of fasciclin II from the extracted supernatant. Likewise, the protein species in lane 4 at 70 kDa is also present in the control lane (not shown) and has not been observed consistently in similar experiments. We interpret

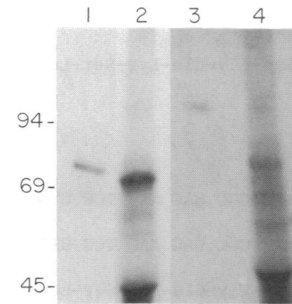


FIG. 3. Fasciclin I behaves as an extrinsic membrane protein, while fasciclin II exhibits characteristics of an integral membrane protein. Membranes prepared by differential centrifugation were labeled with 125 I and lactoperoxidase and subsequently treated with a triethylamine buffer, pH 11.5. The proteins that were released into the supernatant by such treatment (lanes 2 and 4) or were retained with the membranes (lanes 1 and 3) were subjected to immunoprecipitation using mAbs against fasciclin I (lanes 1 and 2) or fasciclin II (lanes 3 and 4). The immunoprecipitates were analyzed on a 10% polyacrylamide gel under reducing conditions.

this to indicate that this species is not a proteolytic fragment of fasciclin II that is released into the supernatant.

In contrast to the results with fasciclin II, approximately 70% of fasciclin I is released from the membrane preparation under identical conditions (Fig. 3, lanes 1 and 2). The remaining 30% of fasciclin I that remains associated with the membranes most likely represents material that is trapped in membrane vesicles and is therefore not released by this treatment. Alternatively, it is possible that a subpopulation of fasciclin I molecules are inserted into the membrane by a phosphatidylinositol or other lipid linkage, as has been observed with a number of other membrane-associated proteins (29), and thus remain associated with the membrane fraction after extraction. This explanation is possible, as the sequence of fasciclin I indicates that the mRNA encodes a signal sequence, but no transmembrane domain (K.Z., L.M., and C.S.G., unpublished data). Thus a mechanism other than a protein insertion through the membrane must be invoked to account for the retention of a portion of fasciclin I with the membranes.

Extraction at high pH typically differentiates extrinsic membrane proteins, which are released into the soluble fraction, from integral membrane proteins, which are retained in the membrane fraction. Such a difference between fasciclin I and fasciclin II may reflect a fundamental difference in their modes of action, despite their similarities in expression and biochemical characteristics. To best study the functional roles of the two proteins, we chose to isolate the genes encoding both molecules.

Protein Sequencing of Fasciclin I and Fasciclin II. The strategy that we adopted to isolate the fasciclin I and fasciclin II genes depended upon obtaining amino acid sequence data from both proteins. Affinity-purified fasciclin I and fasciclin II were further purified by preparative NaDodSO₄/PAGE and electroelution (10). These preparations were subsequently subjected to amino acid sequence analysis on a gas-phase sequencer (30). In the case of fasciclin I, the amino-terminal sequence was determined for 18 amino acids (Fig. 4A). However, fasciclin II was found to be blocked at its amino terminus and thus refractory to sequence analysis. To obtain sequence information from fasciclin II, and to gain additional data from fasciclin I, both proteins were cleaved with CNBr, and the resultant peptides were purified by preparative NaDodSO₄/PAGE and electroelution. One peptide from each preparation was selected for microsequencing (Fig. 4). Although the fasciclin II peptide gave a single amino acid sequence, the fasciclin I peptide gave two different

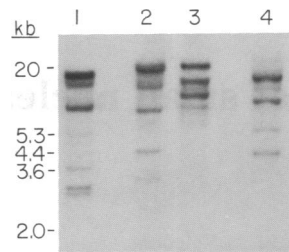


FIG. 5. Southern blot analysis of genomic DNA encoding the fasciclin I and II proteins. Twenty micrograms of grasshopper genomic DNA was digested with *EcoRI* (lanes 1 and 3) or *BamHI* (lanes 2 and 4) and the resultant fragments were separated on an agarose gel. After transfer to GeneScreenPlus, hybridization was performed with ^{32}P -labeled cDNAs encoding fasciclin I (lanes 1 and 2) or fasciclin II (lanes 3 and 4).

gency. The sequence of the 2-kb clone indicated that it encoded the fasciclin II peptide (Fig. 4B). Since clones containing this peptide sequence also are capable of reacting with the antiserum, they are likely to encode part or all of the fasciclin II protein.

Both genes appear to occur in a single copy per haploid genome. As shown in Fig. 5, when grasshopper genomic DNA blots are probed with the fasciclin I and fasciclin II cDNA clones at high stringency, each hybridizes to a small number of bands, consistent with a single copy gene or a small number of very closely related genes. The distinct patterns obtained further emphasize that the two proteins are encoded by different genes.

DISCUSSION

We began these studies with mAbs that recognize two different surface glycoproteins, called fasciclin I and fasciclin II, which are expressed on subsets of axon fascicles in the grasshopper embryo (7). In this paper, the mAbs were used to purify the proteins for further biochemical characterization and microsequencing. This sequence information allowed us to design oligonucleotide probes that we used to clone the genes encoding fasciclin I and fasciclin II.

Fasciclin I and fasciclin II are unique membrane-associated glycoproteins of molecular mass 70 and 95 kDa, respectively; neither protein appears to be covalently associated with any other protein. Fasciclin II appears to be an integral membrane protein, whereas fasciclin I appears to be an extrinsic membrane protein. Both are glycosylated, and both are among a large group of neuronal surface glycoproteins that express the carbohydrate that cross-reacts with antibodies to horseradish peroxidase (28).

The isolation of the genes encoding these two glycoproteins in the grasshopper is a first step toward understanding their structure and function. Both are expressed on a subset of axon fascicles during neuronal development in a manner consistent with a role in cell recognition and growth cone guidance. It will now be of interest to determine the structure of these proteins and search for related proteins. To test the function of these molecules, we would ultimately like to use genetic analysis in *Drosophila*. Given the similarity in the patterns of identified neurons, selective fasciculation, and growth cone guidance between grasshopper and *Drosophila* (3, 31), it should now be possible to use the DNA probes for these genes from grasshopper to search for homologous genes in *Drosophila*.

We thank Zaida Traquina and Violette Paragas for technical assistance. We also thank Hans Acha-Orbea and Ed Fritsch for helpful discussions. This work was supported by American Cancer Society Postdoctoral Fellowships to P.M.S. and A.L.H., a Helen Hay Whitney Postdoctoral Fellowship to K.Z., a National Institutes of Health Medical Scientist Training Program Traineeship to L.M., National Institute of Mental Health Grant MH 23861-13 to G.M., and grants and awards from the National Institutes of Health, McKnight Foundation, and March of Dimes-Birth Defects Foundation to C.S.G.

1. Ghysen, A. & Janson, R. (1980) in *Development and Neurobiology of Drosophila*, eds. Siddiqi, O., Babu, P., Hall, L. & Hall, J. (Plenum, New York), pp. 247-265.
2. Goodman, C. S., Raper, J. A., Ho, R. & Chang, S. (1982) *Symp. Soc. Dev. Biol.* **40**, 275-316.
3. Goodman, C. S., Bastiani, M. J., Doe, C. Q., du Lac, S., Helfand, S. L., Kuwada, K. Y. & Thomas, J. B. (1984) *Science* **225**, 1271-1279.
4. Raper, J. A., Bastiani, M. J. & Goodman, C. S. (1983) *J. Neurosci.* **3**, 31-41.
5. Bastiani, M. J., Raper, J. A. & Goodman, C. S. (1984) *J. Neurosci.* **4**, 2311-2328.
6. Raper, J. A., Bastiani, M. J. & Goodman, C. S. (1984) *J. Neurosci.* **4**, 2329-2345.
7. Bastiani, M. J., Harrelson, A. L., Snow, P. M. & Goodman, C. S. (1987) *Cell* **48**, 745-755.
8. Patel, N. H., Snow, P. M. & Goodman, C. S. (1987) *Cell* **48**, 975-988.
9. Schneider, C., Newman, R. A., Asser, U., Sutherland, D. R. & Greaves, M. F. (1982) *J. Biol. Chem.* **257**, 10766-10769.
10. Hunkapillar, M. W., Lujan, E., Ostrander, F. & Hood, L. E. (1983) *Methods Enzymol.* **91**, 227-236.
11. Konigsberg, W. H. & Henderson, L. (1983) *Methods Enzymol.* **91**, 254-259.
12. Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, L. E. & Weber, P. (1981) *Anal. Biochem.* **118**, 131-137.
13. Haustein, K., Marchalonis, J. J. & Harris, A. W. (1975) *Biochemistry* **14**, 1826-1834.
14. van Agthoven, A., Terhorst, C., Reinherz, E. & Schlossman, S. (1981) *Eur. J. Immunol.* **11**, 18-21.
15. Laemmli, U. K. (1970) *Nature (London)* **227**, 680-685.
16. Morrissey, J. H. (1981) *Anal. Biochem.* **117**, 307-310.
17. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687-701.
18. Mullins, J. I., Brody, D. S., Binari, R. C., Jr., & Cotter, S. M. (1984) *Nature (London)* **308**, 856-858.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294-5299.
20. Gubler, V. & Hoffman, B. J. (1983) *Gene* **25**, 263-269.
21. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194-1198.
22. Wood, W. I., Gitschier, S., Lasky, L. A. & Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1585-1588.
23. Woo, S. L. C. (1979) *Methods Enzymol.* **68**, 389-395.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
25. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
26. Henikoff, S. (1984) *Gene* **28**, 351-359.
27. Jan, L. Y. & Jan, Y. N. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2700-2704.
28. Snow, P. M., Patel, N. H., Harrelson, A. L. & Goodman, C. S. (1988) *J. Neurosci.* **7**, 4137-4144.
29. Cross, G. A. M. (1987) *Cell* **48**, 179-181.
30. Hewick, R. M., Hunkapillar, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem.* **256**, 7990-7997.
31. Thomas, J. B., Bastiani, M. J., Bate, C. M. & Goodman, C. S. (1984) *Nature (London)* **310**, 203-207.